

UNCLASSIFIED

AD

219006

**FOR
MICRO-CARD
CONTROL ONLY**

1 OF 2
Reproduced by

Armed Services Technical Information Agency

ARLINGTON HALL STATION; ARLINGTON 12 VIRGINIA

UNCLASSIFIED

**Best
Available
Copy**

"NOTICE: When Government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the U.S. Government thereby incurs no responsibility, nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications or other data is not to be regarded by implication or otherwise in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related there

AD No 219006

ASTIA FILE COPY

**THE RELATIONSHIP OF METHOD OF INSTRUCTION
TO TRAINEE APTITUDE PATTERN**

FILE COPY

ASVIA

APLINGTON HALL STATION

APLINGTON 12 VIRGINIA

APR. 1958

A RESEARCH PROJECT CONDUCTED UNDER
CONTRACT Nonr 2313(00), OFFICE OF
NAVAL RESEARCH.

JUNE 15, 1958

FC

RICHARDSON, BELLOWS, HENRY & COMPANY, INC.

1 West 57th Street

New York 19, N.Y.

THE RELATIONSHIP OF METHOD OF INSTRUCTION
TO TRAINEE APTITUDE PATTERN

A Research Project Conducted Under
Contract Nonr ~~100000~~ 2313(00)
Office of Naval Research

June 15, 1958

Harold A. Edgerton, Principal Investigator

Richardson, Bellows, Henry & Company, Inc.
One West 57th Street, New York 19, New York

Reproduction in whole or in part is permitted
for any purpose of the United States Government.

ACKNOWLEDGMENTS

Richardson, Bellows, Henry & Company, Inc. takes this opportunity to thank the many persons whose thoughtful cooperation made this study a better experiment.

Among these many persons are the instructors in the AN-P School at NAS Jacksonville, their supervisors, and their advisors.

Special appreciation is expressed to:

DR. D. D. SMITH,
Office of Naval Research, Washington, D. C.,

DR. G. D. MAYO,
Naval Air Technical Training Command, NAS Memphis,

CAPT. C. L. MOORE, JR.,
Commanding Officer, Naval Air Technical Training
Command, NAS Jacksonville,

MR. M. W. ROSS,
Educational Advisor, Naval Air Technical Training
Command, NAS Jacksonville.

Harold A. Edgerton
Principal Investigator

TABLE OF CONTENTS

I	ABSTRACT	1
II	BACKGROUND	3
III	THE EXPERIMENT	
	The Situation	6
	Measurement of Learning	10
	The Instructors	10
	The Aptitude Tests	10
IV	COURSE EXAMINATIONS AS MEASURES OF LEARNING	
	Examination Means	15
	Intercorrelations of Examination Scores	17
V	EQUIVALENCE OF THE EXPERIMENTAL GROUPS	20
VI	VERIFICATION	
	Classroom Observations	24
	Attitude Questionnaire	27
	Walking Chits	29
	Rate Choice	31
	Other Evidences	34
VII	CORRELATION AMONG THE TEST SCORES	36
VIII	RESULTS	
	Correlations of the Tests Scores and Examination Grades	38
	Partial Regression of Aptitude Tests on Examination Grades	40
	Phase I, Unit 4a:	
	Aircraft Familiarization	45
	The Navy Interest Test	47
IX	SUMMARY	50

I

ABSTRACT

To what extent do different teaching methods get better results when different methods are applied to trainees selected because their aptitude patterns are not only appropriate to the subject matter but also to the method in which the subject matter is presented?

From a practical point of view, this study shows that trainees do learn better when the training method is appropriate to their aptitude patterns.

The experiment reported here tested this concept with two methods of teaching. One method of teaching was a strictly rote method. The trainees were required to memorize the material presented and be able to give it back on examinations. The second method asked that, in so far as possible, the instructor present explanations, get the students to ask why, and call for explanations of the various concepts. The experiment was conducted in the AN-P Course, Naval Air Technical Training Center, NAS Jacksonville, a school which is essentially a preparatory school giving refresher training to Marine Corps Airmen just before they are sent to their Class A Schools for training for specific rates. The subject matter of the course includes mathematics, electricity, mechanics, introductory orientation to aviation, and the like. The trainees included were those scheduled to enter the Mechanical Occupations.

Two parts of this course were chosen as the experimental periods of instruction. These were Phase III, Mechanics, Mathematics, and Physics, and Phase I, Unit 4a, Aircraft Familiarization.

Results of the instruction were measured by objective-type examinations given by the school. The same set of examinations was used for all experimental groups.

The results from Phase III were not as clear as those from the other experimental part of the course. The lack of clarity stems from an apparent lack of distinction in the two methods of instruction. The instructors were non-commissioned officers who had been teaching by a more or less rote method and who were not able to make a sharp differentiation between the two methods when asked to use one or the other. However, the differences reflected in the Aircraft Familiarization Unit were clear. It

was found that trainees taught by both the rote method and by the why method required high scores in the Verbal factor. Factor M, Memorization, interfered with success in the why or reasoning method. Factor W, a Word fluency factor, augmented the learning in the rote method.

Vocational interest scores identified with Mechanical Occupations can be used to identify trainees who do well in the rote method while scores for activities more or less opposite to the Mechanical Occupations may be used to identify trainees who do less well when trained by the rote method. It is possible that interest score pattern bears on quality of performance under one instructional method or another.

There is at least a hint in the data collected that the "productive morale" of the trainees is related in a small way to instructional method.

In the case of national emergency, when many men are in training, and where time is very important, it will be possible and feasible to select men for training and to train them with methods appropriate to their aptitude patterns. For example, in the Airman-P School, first get as many men as possible with high V, Verbal factor scores or high OCT scores. Then, split these into two groups for teaching. Those with a low factor M score may be put in one group and taught by the why or explanation method. Those with high factor M scores can be put in a class to be taught by the rote method.

There is still much work to be done on this problem. The experimental work so far has shown that the principle is sound and correct. There has not been enough experimental work done exploring the aptitude patterns demanded by different methods of instruction, nor have the aptitude patterns appropriate to combinations of methods of instruction been studied. Studies need to be made so that, in a case of national emergency, training may be speeded up, and so that, perhaps, a few more men may be available for technical training.

In addition, later on-the-job performance of the experimental trainees needs to be studied relative to the instructional method used and to the corresponding aptitude pattern.

II

BACKGROUND

The research here reported has grown from our observations that, while measured differences among individuals have served a useful function in the selection of persons for training, when the selected persons start their training the training programs make little or no use of the concepts of individual differences. Training method is presumed to be a function of the subject matter and unrelated to the aptitudes of the trainees. Experimentally derived information in this area is basic if we are going to consider a teaching method to fit the aptitude patterns and conceptual habits of the learners, or if we are to select learners by aptitude pattern so as to profit maximally from some distinct method of training.

The major hypothesis of this research is, "Some individuals can learn a subject or body of information better by one method and other individuals can learn that same material better by another method, depending on their pattern of aptitudes."

The hypothesis is not new, but there has been little research to test it and to study its ramifications.

Three prior studies may be cited. Smith¹ showed that intelligence test scores correlated near zero with marks in a high school Industrial Arts course when all grading was done by the instructor, but when projects were evaluated by the pupils themselves, the correlation between marks and intelligence test scores was substantial.

Ellison and Edgerton² showed that college freshman course marks showed substantial correlation with Thurstone's factor V (verbal ability). Marks in biological science showed significant correlation with Thurstone's deductive reasoning factor, while romance language marks showed significant correlation with Thurstone's inductive reasoning factors. The explanation offered was difference in teaching method.

1 Smith, R. E., Unpublished Ph.D. dissertation, 1928, The Ohio State University.

2 Ellison, M. L. and Edgerton, H. A., "The Thurstone Primary Mental Abilities Tests and College Marks", Educational and Psychological Measurement, Vol. I, No. 4, October 1941.

Edgerton³ showed that some trainees learn better when presented with one sequence of materials and others learn better when presented with the same material in a different sequence. This study was conducted in the Aerographer's Mate School, Class A, NATTU, NAS Lakehurst. The course was taught in two patterns:

- A. Insofar as feasible, present the how to do it before the explanation and meteorology.
- B. Insofar as feasible, present the explanation and meteorology before the how to do it.

Trainees who learned better under the how to before explanation arrangement did not have the same pattern of aptitude test score as those who learned better under the explanation before how to arrangement. On the basis of scores on the Chicago Test of Primary Mental Abilities:

Those who were above average in factor V (verbal ability) learned more effectively in both arrangements of the curriculum than did those below average on factor V.

Those who had above average scores for factor N (numerical ability) did better on the more routine aspects of the course when presented how to first.

Those who were above average for factor W (word fluency) did slightly better in the course than those below average on the factor. The differences were greater for those who had the explanation before their how to.

Factor S (spatial discrimination) brings to light the fact that the instruction is routine and procedure oriented, rather than taught in terms of space concepts of weather phenomena.

Those who were above average for factor R (reasoning) did better than those below average in the factor when they had explanation before how to.

Those who were above average for factor M (rote memory) did slightly better than those who were below average on the factor. The difference is greater for those who had the how to before their explanation.

3 Richardson, Bellows, Henry & Company, Inc., "Should Theory Precede or Follow A 'How-To-Do-It' Phase of Training", A research project conducted under contract Nour 1722(00), Office of Naval Research, December 31, 1956.

The present study was designed to explore further the concepts relating individual differences and training methods. In this study, two training methods were selected. These had to be useable in the military training situation and be sufficiently identifiable that the regular instructors could use them. It was also necessary that neither method was likely to result in less well trained personnel than the current instructional methods of the experimental school.

The two selected methods are:

Rote: All materials to be presented and learned by rote methods.

Why: Presenting all materials with as much explanation or why as possible and encouraging the trainees to ask why.

If the hypothesis under investigation in this study is upheld, further research should follow in which similar hypothesis are explored in different kinds of training and educational situations, and using varied kinds of subject matter. These studies would produce the detailed information necessary to give practical implementation to the concepts.

III

THE EXPERIMENT

The Situation

To conduct an experiment to test the ideas set forth, the following factors controlled the selection of a training program from which experimental data could be collected:

1. The training program must contain a substantial segment of material which can be taught either by the rote method or by introducing more explanation or why in presenting the same materials.
2. The school directors and instructors must be willing to cooperate in the experiment, with assurance that it will not cause them to do less effective instruction than by the method they are currently using.
3. Objective form examinations need to be available to measure the quantity or quality of learning.
4. The same set of examinations must be used to measure the learning of all groups of trainees participating in the experiment.
5. A sufficient number of trainees must be entered in the school to make it possible to collect the data economically.

An appropriate setting for the experiment was found in the Naval Air Technical Training Center, NAS Jacksonville, in the Airman-P School. The Airman-P School is essentially a "preparatory" school, giving introductory orientation to aviation, training in elementary manual skills, and refresher training in mathematics, electricity, and mechanics. The training is designed to prepare enlisted men for their training in the Class A Schools which prepare for job performance in a specialized area. The Airman-P School program is shown in the Outline of Training. The materials are presented in an "integrated" rather than in block fashion, i.e., the materials of one phase are interwoven in the instruction to some extent with materials from another phase. If presented in block fashion, the trainees would work full time on one topic, e.g., mathematics, until that was completed before taking up the next topic. An integrated pre-

OUTLINE OF TRAINING

Airman-P School
Naval Air Technical Training Center
NAS Jacksonville

PHASE I - INDOCTRINATION - - - - - 86 hours

- Unit 1 - Introduction to Airman School
(Class P) - - - - - Estimated 8 hours
- Unit 2 - Group testing for placement
into occupational group-
ings - - - - - Estimated 4 hours
- Unit 3 - Aviation rating familiari-
zation - - - - - Estimated 17 hours
- Unit 4 - Airman qualifications - Estimated 51 hours
 - *a. Aircraft familiari-
zation - - - - - 31 hours
 - b. Fire-fighting - - - - - 6 hours
 - c. Survival - - - - - 14 hours
- Unit 5 - Moral guidance - - - - - Estimated 2 hours
- Unit 6 - Guidance (Special
Advisory Program) - - Estimated 4 hours

PHASE II - BASIC MANUAL SKILLS - - - - - 56 hours

- Unit 1 - Safety - - - - - Estimated 12 hours
- Unit 2 - Sheet metal fabri-
cation - - - - - Estimated 12 hours
- Unit 3 - Metal working - - - - - Estimated 16 hours
- Unit 4 - Technical publica-
tions - - - - - Estimated 16 hours

*PHASE III - RELATED SUBJECTS - - - - - 98 hours

- Unit 1 - General Mathematics - - Estimated 40 hours
- Unit 2 - Mechanics - - - - - Estimated 33 hours
- Unit 3 - Electricity - - - - - Estimated 25 hours

*Parts of the course included in this study.

sentation means that trainees receive instruction in each of several topics each day. For example, the first hour might be mathematics, the next two periods mechanics, followed by physics, with the afternoon spent in the shop.

On the basis of review of the curriculum content and from discussion with personnel at the School, it was agreed that the best materials for the purposes of the study would be those presented as Phase III, General Mathematics, Mechanics, and Electricity, and in Unit 4a of Phase I, Aircraft Familiarization. These seemed to offer the best possibility for being taught by both of the experimental methods.

A new class of students enter the School each week, designated by the week of the year they entered, e.g., Class 27 entered in the 27th week of the year, etc. Four classes, 27, 28, 29, and 30, of 1957 were selected for this experiment. The trainees are divided by the school into two groups, MOG (Mechanical Occupational Group) and EOG (Electrical Occupational Group), and are trained in separate sections of the course. The classification of trainees is based on aptitude test scores, work experience, high school training, and interests. The experimental sample included all trainees in the Mechanical Occupations Group (MOG) entering the course during a four-week period. MOG trainees were chosen for the experiment because the subject matter taught them appeared more adaptable to the experiment.

The MOG training leads to schools for the following rates:

PRS Parachute Rigger, Safety	AK Aviation Store- keeper
ADJ Aviation Mechanic: Jet	AG Aerographic Mate
AOU Aviation Ordnance- man: Utility	AOT Aviation Ordnance- man: Turrets
ADH Aviation Mechanic: Helicopter	PH Photographer's Mate
AMH Aviation Metalsmith, Hydraulics	AMS Aviation Metal- smith, Structural
ADE Aviation Mechanic, Reciprocating	PRM Parachute Rigger, Maintenance

Table 1 shows the classes, the dates on which they entered, and the experimental conditions used. The number of classes chosen was determined by the number of trainees available in each class.

TABLE 1

TEACHING METHOD AND SUBJECT MATTER
FOR THE TWO EXPERIMENTAL GROUPS

Group	Class	Beginning Date	Teaching Method Used	Subject Matter
I	27	8 July 57	Rote	Phase III
I	28	15 July 57	Rote	Phase III, I 4a
II	29	22 July 57	Why	Phase III, I 4a
II	30	29 July 57	Why	Phase III

Since seventy-five or more trainees in each class were MOG, two classes would furnish more than 100 cases, and an acceptable number of trainees taught by one of the methods. Two more classes would offer sufficient numbers exposed to the other method. Classes 27 and 28 formed Group I (169 cases) and were taught as completely as possible by rote memory. Classes 29 and 30 formed Group II (158 cases) and were taught by the theory or why method, using reasoning or explanation as an approach to the materials of the course. Each of these individuals had taken of the tests and had completed the course on schedule.

From the point of view of the military composition, the two groups were quite comparable. Table 2 shows the number of MOG trainees in each class according to military composition. Most of the trainees were sent to the school directly from "Boot" camp and had had no other military experience. Class 27 had more six month reserves in it than other classes. It was felt that this did not materially bias the sampling.

TABLE 2

NUMBER OF TRAINEES IN EACH CLASS

Group	Class	Military composition*				Complete Data
		1	2	3	4	
I						
<u>Rote</u>	27	5	70	9	11	83
(N=169)	28	4	81	4	0	86
II						
<u>Why</u>	29	7	63	2	1	71
(N=158)	30	3	80	0	5	87

- *1 Men sent in from a unit.
 2 Recent Boot camp graduates.
 3 Two year reserves on active duty.
 4 Six months reserves on active duty.

The first two classes, Group I, were taught by rote methods, and the last two classes, Group II, were taught by the why methods. This order was chosen because it was believed that it would be easier for the instructors to teach by a purely rote method and that, as they became more acquainted with the experiment, it would be less of a problem for them to shift to the why method than to reverse the procedure. Also, the experimenter held the personal belief that the more effective instructors would prefer to put more why into their teaching as they began using more of that particular approach.

Measurement of Learning

The effect of instruction was measured by subject matter examinations prepared by the school. All four experimental classes were given the same examinations to provide common standards by which their learning could be measured and compared. The examinations were multiple choice type. Each of the questions had been used previously and had been checked by item analysis for difficulty level, effectiveness of the distractors, and discrimination of high and low total scores. In order to give some security to the examinations, and minimize informal passing of examination information from one class to the next, the order of presentation of the questions in each examination for the four classes was varied from one class to the next.

The Instructors

There were eleven instructors in Phase III during the experimental period. All instructors taught their subject matter in the course using the rote method in Classes 27 and 28 and then used the why method for Classes 29 and 30. The only exceptions to this were due to normal attrition and leave. This had the effect of removing "instructor-method" interaction from the experimental design. The instructors were Marine Corps NCO's who were trained and experienced in the various aviation rates.

The Aptitude Tests

To obtain patterns of aptitudes which might have significance for this study, it was desirable to have measures of a number of relatively independent aptitudes or aptitude factors which presumably might be related to habits of learning. In addition each measure need to be appropriate to the intellectual and educational level of the experimental subjects.

The Chicago Test of Primary Mental Abilities was selected because it met the specifications and because it had indicated some usefulness for such a purpose in earlier studies. This timed test was administered to all trainees included in the study. For Class 27, the test was administered during the first week of instruction and for the other three classes, the test was administered during their orientation week prior to class attendance. Scores are produced for each of six Primary Mental Factors, defined as follows:

Factor N - The Number factor involves the ability to do numerical calculations rapidly and accurately. It is not dependent upon the reasoning factors in problem-solving but seems restricted to the simpler processes, such as addition and multiplication.

Factor V - The Verbal factor is found in tests involving verbal comprehension, for example, tests of vocabulary, opposites, synonyms, completion tests, and various reading comprehension tests.

Factor S - The Space factor involves any task in which the subject manipulates an object imaginably in space.

Factor W - The Word Fluency factor is called upon whenever the subject is asked to think of isolated words at a rapid rate. It can be exhibited in such tests as anagrams, rhyming, and recalling words with a given initial letter, prefix, or suffix.

Factor R - The Reasoning factor can be found in tasks requiring the subject to discover a rule or principle involved in a series or groups of letters. In experimental studies, two separate Reasoning factors have been indicated, perhaps Induction or Deduction. The present reasoning tests are more heavily saturated with the Inductive factor, but the factor here is called Reasoning R.

Factor M - The rote Memory factor depends upon the ability to memorize quickly.

The scores on these six factors are relatively independent. Each factor measures some component of what has been recognized as "general intelligence." Hence, it is reasonable to believe that the factor scores will identify persons having differing patterns of ability. Persons with some patterns of abilities will learn better under one method of presenting the subject matter and others, who have some other pattern of scores, will learn better under other methods of presentation.

In addition to the Chicago Test of Primary Mental Abilities, certain test scores were available for all of the trainees. Since these tests were regularly administered and used in the classification and placement of trainees, it appeared pertinent to include such test scores as possible sources of information regarding aptitude patterns. Apparently, the only effect pertinent to this study in using these tests as indicated is one of reducing the variability of the test scores of those trained as MOG's relative to Navy enlisted personnel as a whole. Data from three such tests were available from the trainee records; the General Classification Test, the Navy Mechanical Comprehension Test, and the Navy Vocational Interest Inventory.

The General Classification Test, a test of the individual's ability to learn, is currently used by the Marine Corps for purposes of recruit classification. It was developed by the Army and adopted for Marine Corps use about 1948. It is a four-part, multiple-choice type test with questions in each part arranged in ascending order of difficulty. Separate scores were available for each of the three parts or sub-tests, as well as a total score for the entire test.

The sub-tests are described briefly below:

Reading and Vocabulary - This is a 53-item sub-test requiring twenty-five minutes. It consists of a series of paragraphs with certain words in each underlined. The individual is presented with multiple-choice type of questions which ask him to identify the meaning of the underlined word and additional multiple-choice questions designed to test his understanding of the material presented.

Arithmetic Computation - This is a 52-item sub-test requiring fifteen minutes. It measures the subject's ability to solve arithmetic problems by reasoning.

Pattern Analysis - This is a 50-item sub-test requiring twenty minutes. Items in this sub-test are a series of unfolded patterns together with a corresponding series of folded figures. The individual is presented with both the pattern and the figure and is required to identify portions of the figure which match those appearing on the pattern. It tests his ability to manipulate patterns mentally.

Total Score - The total score on the General Classification Test was also available, in addition to the scores for the sub-tests which comprise it. As a whole, the test is designed to determine the extent of a trainee's ability to learn.

The Navy Mechanical Comprehension Test, administered as part of the Navy Basic Test Battery, measures the trainee's mechanical comprehension.

The Navy Vocational Interest Inventory was administered to all trainees during Phase I, Unit 2, of the School training. This had been scored for only four rates. The scores on these are used as aids in classifying trainees for training for different rates. Scoring keys for the remaining twelve rates were made available and scores were obtained for this study. The Navy Vocational Interest Inventory uses the expressed preferences for tasks and activities of an individual to show the degree to which that person's preferences match the tasks and activities characteristic of various Navy

specialties. The Inventory included 570 items arranged in groups of threes, making a total of 190 triads. The test task is to indicate which of the activities shown in each triad is most preferred. It is untimed.

In addition, scores were obtained for each of nine general or homogeneous keys. These keys were constructed by combining or amalgamating keys for rates which showed a high degree of similarity.

These homogeneous keys may be characterized as follows:

H1 Mechanical - The items in this key relate primarily to mechanics, machine operation and design, or to home repairs of mechanical and simple electrical gadgets.

H2 Medical Service - The items in this key express interests in medical and hospital service activities and occupations, or in doing medical, biological, or chemical research.

H3 Clerical - Interests in general office clerical work, office machine operation, bookkeeping and accounting, and office management are indicated by the items scored on this key.

H4 Electronic - These items deal with the maintenance, operation and building of radio and other electronic equipment and with the repair and construction of electrical systems and devices.

H5 Food Service - These items are concerned almost completely with interests in the preparation of food and in menu planning.

H6 Woodworking - The content of the major cluster apparent in this key deals with carpentry and furniture-making. The remaining items (plus some in the above cluster) seem quite unrelated, although each selected item involves the rejection of an alternative that deals with electrical-electronic or with medical-chemical interests.

H7 Verbal Activities - Two clusters of content are indicated by these items. The largest deals with a variety of verbal activities while the other indicates interests in aesthetics. A few of the items express an interest in people; others seem related only in that they seem socially accepted, "highly-thought-of" activities.

H8 Clean Hands Activities - There seems to be no easily interpretable common theme indicated by these items, although it appears that high scores on this key reflect preference for "clean hands" kinds of activities.

H9 Physical Effort - The major cluster of items in this key reflects interests in athletics and other outdoor activities. A second set of interests deals with unskilled manual jobs and home repairs. The remainder seem to indicate a sort of compulsiveness or neatness about simple manual tasks. The alternatives to the keyed responses reflect aversions for feminine, indoor, verbal and responsible complex activities.

IV

COURSE EXAMINATIONS AS MEASURES OF LEARNING

The school examinations furnish the basic measures by which the pertinent hypotheses may be tested. All trainees in the experiment were given the same examinations, whether they were taught by rote method or by the why method. The examinations were of a four-answer multiple-choice type. The examination questions had been drawn from the course materials, and had been subjected to item analysis to be sure that each question did differentiate between trainees who got high scores on the examination as a whole and those who got low scores. The item analysis was also used to be sure that each answer alternative was useful, and that the question was of appropriate difficulty. Since these examinations had been well constructed and appeared to sample the course content well, they were judged to be satisfactory measures of learning for the purpose of the experiment.

Six examinations were used to cover Phase III. These were:

Mathematics I	Mechanics I	Electricity I
Mathematics II	Mechanics II	Electricity II

One examination covered Aircraft Familiarization (Phase I, Unit 4a). This was analyzed only for Classes 28 and 29. Examinations given over other phases of the course (Manual Skills, Fire-Fighting, and Technical Publications) were analyzed, but were not considered as part of the experiment since the teaching methods were not controlled in those segments of the course.

Examination Means

Table 3 shows how the two groups differed on the course examination scores. The trainees taught by the why method (Group II) showed on the average a performance superior to the trainees taught by the rote method (Group I) in all examinations of Phase III except one, Electricity I. This difference between the two groups was significant at the 5% level of confidence for the examinations covering Mathematics II and Electricity II, and at the 1% level for examinations in Mathematics I and Mechanics I. The differences in course achievement for Groups I and II are greater than might be expected on the basis of differences in aptitude test scores. The fact that Group II took the same examinations as did Group I, but at a later date, may have had some

influence on the observed differences. However, these findings do tend to suggest that trainees learn course materials better if they are taught by the why method.

TABLE 3
COMPARISON OF GROUPS I AND II ON COURSE EXAMINATIONS

(Rote: Group I, N = 169)

(Why: Group II, N = 158)

Course Examination or Performance Measure	Mean		Standard Deviation		Critical Ratio
	I	II	I	II	
Mathematics I	76.51	80.96	13.87	11.70	-3.16 **
Mathematics II	74.44	78.74	18.28	16.66	-2.23 *
Mechanics I	76.33	79.37	10.97	9.53	-2.69 **
Mechanics II	78.46	80.13	10.10	9.17	-1.56
Electricity I	80.83	80.51	13.13	13.07	0.22
Electricity II	73.90	76.59	11.39	8.93	-2.37 *
Manual Skills	73.55	74.63	8.99	7.63	-1.17
Fire-fighting	73.61	75.01	9.22	9.37	-1.36
Aircraft					
Familiarization	77.40	80.32	10.68	10.20	-2.54 *
Technical Publications	74.80	78.36	9.51	8.09	-3.67 **
Final Average	75.74	77.92	9.68	9.05	-2.10 *
Performance Achievement Quotient (PAQ)	53.72	55.89	10.19	10.48	-1.90

* Critical Ratio of difference between group means significant at the 5 per cent level of confidence.

** Significant at the 1 per cent level.

As a matter of passing interest, although no control was made over the teaching methods, Group II trainees again were superior to Group I trainees on examinations in Manual Skills, Fire Fighting, and Technical Publications. It could be argued, without benefit of definitive evidence, that Group II trainees had the benefit of "leakage" of examination questions from the immediately preceding Group I trainees.

By averaging all of the examination grades within the course, a "final average" grade was obtained. As one would expect, the Group II trainees, taught by the why method, were found to show a higher final average than trainees of Group I, taught by rote

methods. The difference for the final average was significant at the 5% level of confidence.

One further evidence of performance in the course was a comparison of the Performance Achievement Quotient (PAQ) for Groups I and II, shown in Table 3 above. The PAQ was produced statistically by eliminating from the final average grade the variance associated with the score on the General Classification Test. It is the best answer to the question, "Intelligence equal, how well did the trainees get along in the course?" It will be noticed from Table 4 that, although Group II trainees (why) were superior to Group I trainees (rote), this tendency was not strong enough to be considered significant. The Critical Ratio obtained was 1.90, and to be significant at the 5% level of confidence, this ratio would have to be 1.96 or greater.

In summary, in performance on course examinations, there were strong indications that the Group II trainees, taught by the why method, were superior to the Group I trainees taught by rote method. By and large, this difference is great enough to be considered a significant finding. However, the real purpose of this study is to see if trainees with one pattern of aptitudes are the ones who do better by the why method and others having a different pattern of aptitudes would do better when the rote method is used.

Intercorrelations of Examination Scores

Table 4 shows the intercorrelations among the several examinations in the course. The average of the intercorrelations of the course grades for Group I were higher ($\bar{r} = .583$) than those for Group II ($\bar{r} = .496$). The correlation between the Group I coefficients and the Group II coefficients (78 pairs of coefficients) was .68. This probably reflects the fact that the examinations as constituted were more adapted to rote learning than to the why method.

Inspection of the two sets of intercorrelations suggests that a single factor could account for practically all of the covariance among the course marks. Using Thurstone's centroid method of factor analysis, the first centroid factor was obtained from the correlations among all of the examinations. The factor loadings for each of the examinations are shown in Table 4 in Row F for Group I and in Column F for Group II. These factor loadings are all high and the similarities of the loadings and the pattern of loadings for the various examinations is noticeable.

A similar set of factor loadings was computed, designated as F_1 in Table 4, based on the intercorrelations of only those examinations within Phase III. The loadings are, of course, slightly higher than those computed on the basis of all examinations, but do follow the same patterning.

The concept that the examination grades represent only a single factor is supported by the magnitude of the first factor (f_1) residuals. Only two of the 15 residuals for Group I differed significantly (5% level) from zero. This was also the case for Group II.

TABLE 4

INTERCORRELATIONS* BETWEEN COURSE EXAMINATION SCORES

(Note: Group I below the diagonal, N = 169)
 (Why: Group II above the diagonal, N = 158)

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	F	F ₁
Math I (1)	—	701	501	522	454	386	438	490	353	466	700	-725	544	738	732
Math II (2)	775	—	643	606	551	506	459	440	439	564	802	-819	602	824	832
Mechanics I (3)	621	608	—	649	547	535	427	469	541	495	702	-742	580	799	790
Mechanics II (4)	653	636	698	—	536	513	496	490	483	466	706	-761	576	795	779
Electricity I (5)	560	610	721	658	—	397	367	296	426	511	570	-673	569	694	681
Electricity II (6)	505	505	600	529	591	—	386	403	443	422	577	-619	491	684	644
Manual Skills (7)	495	536	504	524	567	543	—	460	397	371	607	-621	433	664	—
Firefighting (8)	412	452	613	533	594	524	511	—	486	317	562	-578	477	665	—
Aircraft Famill. (9)	472	446	636	617	624	442	509	519	—	479	504	-566	590	682	—
Tech. Publicat. (10)	525	456	562	610	587	433	377	509	568	—	597	-603	536	698	—
Final Average (11)	774	823	770	797	739	661	700	622	663	619	—	-821	641	900	—
Final Rank (12)	-794	-826	-793	-801	-786	-699	-714	-674	-693	-640	-921	—	-664	—	—
Perf. Ach. Quot. (13)	562	624	618	624	618	551	551	435	577	529	707	-720	—	—	—
F (Factor loadings: All Exams)	783	789	845	836	836	730	722	725	749	720	941	—	—	—	—
F ₁ (Factor loadings: Phase III Exams)	814	818	831	810	808	697	—	—	—	—	—	—	—	—	—

* Decimal points omitted.

EQUIVALENCE OF THE EXPERIMENTAL GROUPS

A necessary condition to an experiment of this type is that the groups of trainees whose performance is to be compared must be equivalent and comparable at the beginning of the experiment. The equivalence must be in terms of those characteristics or measures of the groups which presumably could be affected by the nature of the experimental process.

The two experimental groups were compared for equivalence in terms of tests administered prior to or at the beginning of this training. Means of the test scores and sub-test scores of Group I (Classes 27 and 28) have been compared* with those of Group II (Classes 29 and 30).

Table 5 shows the mean score for the six factors of the Chicago Test of Primary Mental Abilities Test for Group I and for Group II.

TABLE 5
COMPARISON OF GROUPS I AND II
ON THE CHICAGO PRIMARY MENTAL ABILITIES TEST

(Note: Group I, N = 169)
(Why: Group II, N = 158)

Factors	Mean		Standard Deviation		Critical Ratio
	I	II	I	II	
N-Numbers	54.80	56.84	16.94	18.56	-1.04
V-Verbal	62.26	63.87	16.83	16.02	-0.96
S-Space	56.33	57.66	20.11	21.93	-0.57
W-Word Fluency	53.91	51.90	15.41	14.06	1.23
R-Reasoning	28.30	29.02	8.66	8.86	-0.75
M-Rote Memory	8.30	8.94	3.87	4.48	-1.38

* In each case the comparison was made in terms of the critical ratio. This statistic was obtained by dividing the difference in means for the two groups by the standard error of that difference. Critical ratios between 1.96 and 2.58 indicate differences of such magnitude that the two means could be drawn from the same universe less than once in 20 times, and are interpreted as of "moderate" significance. Values of 2.58 or greater may occur in the same universe less than once in 100 times. Since the probability that such differences could arise by chance less than once in 100 times, these differences are accepted as stable or significant. Positive values of a critical ratio show that the Group I mean is greater than the Group II mean, while negative values indicate the reverse.

None of the critical ratios comparing the means of the two groups were as great as 1.96, the 5% level of significance. This permits the assumption that the two groups are random samples of the same universe, and hence comparable.

In Table 6 are shown the means of Group I and Group II for the Marine Corps Aptitude Tests.

TABLE 6

COMPARISON OF GROUPS I AND II
ON THE MARINE CORPS TESTS

(Note: Group I, N = 169)
(Why: Group II, N = 158)

Test	Mean		Standard Deviation		Critical Ratio
	I	II	I	II	
GCT Sub-Tests:					
Reading and Vocabulary	53.75	55.45	8.39	7.35	-1.95
Arithmetic Computation	50.43	52.03	7.42	6.81	-2.03 *
Arithmetic Reasoning	50.19	51.78	7.62	6.40	-2.04 *
Pattern Analysis	55.22	55.16	7.96	9.16	0.06
Total Score	53.33	54.56	6.61	6.95	-1.76
Navy Mechanical Comprehension	52.09	53.27	8.09	7.87	-1.33

* Critical Ratio is significant at the 5% level.

The means are in terms of Navy standard scores (for a standard recruit population, $M = 50$ and $\sigma = 10$). Two of the sub-tests, Arithmetic Computation and Arithmetic Reasoning, showed differences significant at the 5% level of confidence. The means of the Reading and Vocabulary Test for the two groups differed at almost the 5% level. In all three cases, the mean test scores for Group II were higher than those of Group I. These differences, however, are small enough that, considered along with those shown in Table 5, they permit the conclusion that the trainees of Groups I and II may be considered random samples of the same universe of trainees.

The means of the scores on the Navy Interest Inventory for Groups I and II for each of the twenty-five keys are shown in Table 7. The means are expressed in terms of raw scores. The difference in means of the two groups were significant at the 5% level of confidence in five of the twenty-five keys: DC-Damage Controlman, EM-Electrician's Mate, HA-Electronics, H6-Woodworking, and ET-Electronics Technician. The difference in means was significant near the 1% level for Electronics Technician. None of these differences, except perhaps H6, relate to rates pertinent to the Mechanical Occupational Groups (MOG) trainees.

The differences between Groups I and II, as shown in Tables 5, 6, and 7 suggest that the two groups may be considered approximately equivalent, particularly when considering them in terms of the aptitude patterns reflected by the Chicago Test of Primary Mental Abilities. Since the condition of comparability of the two experimental groups has been met, it is appropriate to examine the further data relative to the major hypothesis of the study.

TABLE 7

COMPARISON OF GROUPS I AND II ON 25 KEYS
OF THE NAVY VOCATIONAL INTEREST INVENTORY

(Note: Group I, N = 169)

(Why: Group II, N = 158)

Vocational Key	Mean		Standard Deviation		Critical Ratio
	I	II	I	II	
(AT) Aviation Technician (Electronics)	3.21	8.29	9.44	9.96	-0.37
(AD) Aviation Mate (Mechanical)	16.93	16.65	7.96	7.23	0.35
(CP) Construction Apprentice	- 8.10	- 8.09	9.44	8.36	-0.01
(AM) Aviation Structural Machinist	13.67	11.97	10.40	9.20	0.16
(AO) Aviation Ordnanceman	4.26	4.50	6.43	6.75	-0.32
(BM) Boatswain's Mate	- 1.65	- 4.55	7.77	7.78	1.35
(CS) Commissaryman	-13.25	-14.55	8.48	9.40	1.31
(CT) Communications Technician	- 6.03	- 7.21	10.39	8.21	1.14
(DC) Damage Controlman	1.60	- 1.07	10.17	9.24	2.49 *
(EM) Electrician's Mate	11.13	14.50	13.18	12.38	-2.39 *
(ET) Electronics Technician	0.54	3.99	11.26	12.92	-2.57 *
(FC) Fire Controlman	5.51	7.60	10.07	10.96	-1.80
(GM) Gunner's Mate	20.00	19.18	11.91	10.65	0.66
(HM) Hospital Corpsman	- 1.54	- 1.39	9.50	10.26	-0.14
(RD) Radarman	- 1.12	- 0.44	8.05	8.70	-0.74
(SK) Storekeeper	8.51	8.58	3.18	2.86	-0.20
(H1) Mechanical	15.17	15.24	5.03	5.23	-0.23
(H2) Medical Service	3.98	4.11	3.67	3.69	-0.31
(H3) Clerical	3.43	3.41	2.86	3.01	0.07
(H4) Electronics	9.28	10.47	4.76	4.73	-2.27 *
(H5) Food Service	3.20	2.87	2.84	2.93	1.04
(H6) Woodworking	7.53	6.36	4.20	4.11	2.54 *
(H7) Verbal Activities	3.55	3.77	2.89	3.11	-0.64
(H8) Clean Hands Activities	4.32	4.47	2.25	2.56	-0.57
(H9) Physical Effort	10.11	9.97	2.58	2.64	0.49

* Critical Ratio significant at the 5 per cent level of confidence.

VI

VERIFICATION

There is need for assurance that two methods of instruction were used and that they were really different. Several ways were tried in order to obtain this assurance.

1. The methods of instruction wanted were discussed with the instructors.
2. Class room procedures were observed frequently.
3. Trainees were asked to answer some attitude questions regarding their training in Phase III of the course.
4. Other evidences were sought in terms of trainee rate preferences.
5. Other evidences were also sought in their record of walking chits.

Before starting the collection of data, a meeting was held with the instructors of the phase involved in which the experiment was described and discussed. The instructors' role was indicated by describing the teaching method they were to use. At this first meeting instructors were asked to use the rote method of teaching. The presentation consisted of telling about the experiment, how it was to be conducted, possible implications for practical use, examples of instructional procedure, and discussion of how to deal with trainees' questions. Questions raised by the instructors and by their supervisors were discussed and answered.

Just prior to the beginning of Classes 29 and 30, (group II) the instructors were called into a second meeting to discuss teaching by the why method. Substantially the same instructors taught the subject matter for both methods of instruction. Discussion covered instructional techniques, how to get trainees to raise appropriate questions, how to look for explanations, and the why of any principles, facts, problems, or procedures.

Classroom Observations

The observer reported that he frequently had informal visits

with one or two instructors discussing the experiment and their part in it.

The trainees were not informed that there was any experimental work going on. All they knew was that they had been required to take the Primary Mental Abilities Tests, and that in some of the classes a civilian observer showed up frequently. This was not an unusual event, since civilian supervisors of instruction frequently visit the classes.

During the entire experimental period, an observer visited classes in Phase III and in Phase I, 4a each day. The observer kept a record of when and how long he visited each class, his estimate of the proportion of time he observed each method of instruction, and, for each visit, how many times why or explanation occurred. The Class Visit Report form is shown on the following page. After some observations, it became clear that a record of the number of why's or explanations was the simplest and most objective observation to make within the amount of time each class could be visited. Such behavior could be observed and recorded. Table 8 shows that for Classes 27 and 28 (Group I) the number of why's or explanations per hour was respectively 1.3 and 1.4 during the observations of Phase III instructions.

TABLE 8

SUMMARY OF CLASS VISITS TO OBSERVE TEACHING METHODS

Class & Phase	Teaching Method	Total Observation	No. of Sessions Visited	No. of Why's Observed	Why's Per Hour
27, Phase III	Rote	21.3 hrs.	90	27	1.3
28, Phase III	Rote	18.1 hrs.	81	25	1.4
29, Phase III	Why	19.4 hrs.	89	79	4.1
30, Phase III	Why	14.3 hrs.	66	62	4.4
28, Phase I, 4a	Rote	3.5 hrs.	12	1	.3
29, Phase I, 4a	Why	3.0 hrs.	12	16	5.3

These answers are based on 90 and 81 observations totalling 21.3 and 18.1 hours, respectively, spent observing in class rooms. For the why method, the numbers of why's showed an increase to a rate of 4.1 and 4.4 per hour for classes 29 and 30 respectively. A greater difference is shown between the two teaching methods in Phase I, 4a, Aircraft Familiarization. Why's per hour were 0.3 and 5.3 for the rote and why methods, respectively. These evidences support the presumption that there were differences in the teaching procedure for the two experimental groups, and that the differences for Aircraft Familiarization, Phase I, Unit 4a, were more distinct than for Phase III.

CLASS VISIT REPORT

KEH Project 741

Class: 27 28 29 30

Date _____

Phase: III IVa

Name of Instructor _____

Unit No. _____

Lesson No. _____

Time visit began _____ Time ended _____

Duration _____ minutes

1. Class procedures observed (Estimate % of time for each during visit).

_____ Lecture

_____ Discussion

_____ Demonstration

_____ Oral Questioning by Instructor

_____ Oral Questioning by Students

_____ Other (Specify) _____

2. How many instances of "why" or "explanation" or "reason for" were observed? _____

Characterize each instance briefly.

Several less direct evidences of trainee reaction to the two instructional methods were sought. These included:

1. A trainee attitude questionnaire administered upon completion of Phase III.
2. Collection of data regarding the numbers of "walking chits" issued to members of the experimental groups.
3. Collection of honor class ratings.
4. Examination of the rate choices expressed by the trainees.

These evidences, particularly the last three named, could be thought of as expressing the attitude and morale of the trainees to the training situation. Collection of such data stemmed from the development of the Employee Relations Indicator.*

Attitude Questionnaire

It was thought that differences in the methods of instruction for the two groups might be reflected in the trainees' attitudes toward certain aspects of their training. Hence, an attitude questionnaire was constructed and administered to the trainees during the last day of instruction in Phase III. The 12 questions used in the questionnaire are shown in Table 9. The directions for the questionnaire indicated that the questions referred to only Phase III. Trainees were to indicate how they felt about each of the questions or statements by circling either the word "agree" or "disagree."

Of the 12 statements, only four were presumed to be "loaded" for the purpose of this study. These were:

- No. 2 Phase III had too much memorizing.
- No. 6 There were not enough explanations in the teaching.
- No. 7 There was too much theory in the course.
- No. 9 There was too much repetition of each thing to be learned.

* Merrihue, Willard V. and Katzell, Raymond A. "EMI--Yardstick of Employee Relations," Harvard Business Review, November-December, 1955, p. 91.

TABLE 9

COMPARISON OF GROUPS I AND II ON
THE ATTITUDE QUESTIONNAIRE(Note: Group I, N = 170)
(Why: Group II, N = 162)

Attitude Item	Per Cent Marking "Agree"		Critical Ratio *
	Group I	Group II	
1. The material in Phase III was too difficult.	4	5	-.44
2. Phase III had too much memorizing.	23	24	.83
3. Too much lecture was used in teaching.	15	12	.80
4. We were not given enough practical training.	44	39	.93
5. There were too many examinations.	5	3	.83
6. There were not enough explanations in the teaching.	24	27	-.63
7. There was too much theory in the course.	24	25	-.21
8. The material of the course was presented too fast.	52	52	0
9. There was too much repetition of each thing to be learned.	11	15	-1.08
10. The examinations were too hard.	7	2	2.23
11. Too much discussion was used in teaching.	6	5	.40
12. The examinations did not cover what was taught.	7	6	.37

* Positive signs show that Group I had a greater per cent marked "agree" than did Group II.

A priori, the differences in per cents marking Agree; Group I minus II, were expected to be positive for statements 2, 6, and 9, and negative for 7. Table 9 also shows the per cents of trainees in each of the two experimental groups who indicated agreement with each statement. The last column in the table, Critical Ratio, expresses the difference between the two sets of responses. Where the critical ratios have negative signs, it shows that Group II had a greater per cent who marked "agree" than did Group I.

None of the four "loaded" questions showed a difference of any statistical significance. In the entire questionnaire only Item 10, "The examinations were too hard," showed a difference significant at the 5% level of confidence. However, only a very small per cent of either group agreed with the statement.

On the basis of the foregoing evidence, it seems clear that two teaching methods did not make any differential impact on the trainees as far as Phase III was concerned.

It is possible that attitude questions could have elicited some differences had the questions been empirically pre-selected for their sensitivity to the learning situation. This step was not feasible in the study. These results may also reflect the hypothesis that there was little difference in teaching method between Groups I and II.

It is perhaps likely that the attitude questionnaire would reflect differences only if one teaching method were used exclusively. That is, most of the attitude statements were phrased in comparative terms; e.g., "Too much" in a situation where comparison was not easily possible. The statements would have meaning therefore only in terms of intensity of reaction, an intensity which was apparently lacking.

Walking Chits

Following the development of ERI, Employee Relations Indicator, it seemed important to look for behavioral symptoms of attitudes. For the experimental samples of trainees, these behavioral symptoms seem to be available in rate choice, walking chits, and possibly in honor company rating.

Table 10 shows a walking chit analysis for Class 27 of the Group I: Note trainees. This is presented to show the number of chits issued for one class as a whole; and various kinds of reasons for which chits were issued. Some men may have had more than one chit for the same reason.

TABLE 10
WALKING CHIT ANALYSIS
(for Class 27)

Reason for Chit	Number of Chits
Sickbay	17
Personnel (allotments and pay)	7
Dental	4
To see doctor *	4
Counseling	4
Shots	2
Academic conference	1
Check on guard	1
Chaplain	1
To Security for traffic ticket	1
To see First Sergeant	1
Business in Jacksonville	1
Total	44

* This is different from normal sickbay in that the trainee is told to report back.

Following this an attempt was made to classify the walking chits according to four headings.

1. Sickbay (medical attention of any sort which requires that a man ask for a walking chit).
2. Other medical chits.
3. Chits for which the man did not apply, those which he was "sent for."
4. Chits for any other reason.

The per cents of persons who had chits for each of the four reasons were 20 per cent, 2 per cent, 12 per cent and 2 per cent, respectively, for the four classes. Because of this distribution, it was decided not to subdivide the chits by reasons.

Table 11 shows the point-biserial correlations between the final average of grades earned in the course examinations, and having or not having walking chits. None of the obtained coefficients was significantly different from zero.

TABLE 11

CORRELATIONS BETWEEN CHITS AND
FINAL AVERAGE IN COURSE EXAMINATIONS

(Note: Group I, N = 169)

(Why: Group II, N = 158)

Type of Chit	Group I	Group II
Sickbay (Voluntary)	-.050	.030
Other Medical	-.120	.060
Not Voluntary	.040	-.120
Total Number of Chits	-.080	-.020

For those trainees for whom there were rate choice data as well (N = 155, Group I and N = 143, Group II), 32 per cent in Group I had one or more walking chits, while 24 per cent in Group II had walking chits. The critical ratio for this difference is 1.49. If this could be taken as evidence of attitude (e.g., evasion of class attendance), the why presentation would seem to be better liked than the rote. This evidence may be related to the fact shown later that Group II (why) average examination marks were higher than for Group I.

Rate Choice

Among the evidences of attitude toward training is the rate choice. Each trainee early in the training program records his choice of rates he would like to try for. He does this three more times during the course of the six weeks training, after receiving some instruction including general verbal information about the rates and sample experience of the kinds of skills required. The final evidence is the rate assigned.

The rates themselves are distinct and discrete categories and form no necessary scale or system of values in or of themselves. It was thought that the rates must have meaning to the trainees other than just the kind of vocational activity. After discussing this with several of the instructors, the investigator asked that the several rates available to the MDG trainees be rated according to their prestige, their potential for advancement in the service, for their reputed difficulty, and for their x value in civilian living. The list of twelve rates which were open to the mechanical occupations groups of trainees were listed. Two groups of raters were chosen; a group of instructors and a group of trainees. Each rater was asked to pick the three rates which had the highest prestige and the three having the lowest prestige. They were asked to do the same for advancement in the service, degree of difficulty, and in terms of good preparation for civilian jobs. A score for

each of the factors was obtained by giving one point for each time the rate was nominated as a high, a minus one for each time it was nominated low. The score was the number of high nominations minus low nominations. The scores for the four factors were correlated as shown in Table 12.

TABLE 12

CORRELATIONS OF FOUR RATINGS FOR THE RATES

	1	2	3	4
1. Prestige	—	.91	.81	.35
2. Advancement in Service	.91	—	.86	.62
3. Difficulty	.81	.86	—	.52
4. Civilian Value	.35	.62	.52	—

The correlations among the first three factors; prestige, opportunity for advancement in the service, and degree of difficulty were all very high. The lowest of these correlations was .81. The correlations of these three factors with good preparation for civilian jobs were lower. The scoring values then for the rates are based on the aggregate of the scores for the first three factors and are as follows:

- 0 - AOU Aviation Ordnanceman: Utility
- 2 - ADM Aviation Mechanic: Reciprocating
- 3 - AK Aviation Storkeeper
- 4 - AOT Aviation Ordnanceman: Turrets
- 4 - PM Parachute Rigger: Maintenance
- 5 - AMH Aviation Metalsmith: Hydraulics
- 6 - ADM Aviation Mechanic: Helicopter
- 6 - AMS Aviation Metalsmith: Structural
- 6 - PH Photographer's Mate
- 6 - PRS Parachute Rigger: Safety
- 7 - AG Aerographer's Mate
- 9 - ADJ Aviation Mechanic: Jet

The means and standard deviation of the first, second, etc. times of rate choices were expressed as shown in Table 13 for those trainees who were in a position to make a choice. It may be noted here that some trainees were sent to the school with a predesignated rate so that they had no choice.

TABLE 13

MEANS AND STANDARD DEVIATIONS OF
RATE CHOICES

Time of Choosing	Mean	Standard Deviation
First	7.16	1.29
Second	6.27	2.87
Third	6.17	2.99
Fourth	5.56	3.31
Assignment	4.49	3.38

There is a distinct decrease in the mean values of the choices from the first to the fourth time a choice was expressed, with a correlative increase in variability, perhaps reflecting an increasing acquaintance with reality.

The mean of last expressed rate choice values for Groups I and II were 5.41 and 5.54 respectively. This difference is not significant.

Table 14 shows the intercorrelations of the choices. The correlations between adjacent choices are higher than between choices further apart in time, except for the first and second choosings.

TABLE 14

INTERCORRELATIONS OF VALUES FOR RATE CHOICES

	First	Second	Third	Fourth	Assgd.
First Choosing	—	.19	.20	.43	.20
Second Choosing	.19	—	.64	.56	.41
Third Choosing	.20	.64	—	.74	.53
Fourth Choosing	.43	.56	.74	—	.75
Assigned Rate	.20	.41	.53	.75	—

These evidences suggest the notion that those trainees whose rate choice changes most, particularly in a downward direction, may be less satisfied with their experience in the school. On the basis of these evidences, two measures of rate choice were selected: the last expressed choice, and the value of the last rate choice minus the value for the first.

The per cents in each of Groups I and II for which the difference between the last and first expressed rate choice was positive, zero, and negative are shown in Table 15. There is only a small difference in the per cents of Groups I and II who showed a posi-

tive difference, e.g., the last expressed choice had a higher value than did their first expressed choice.

TABLE 15

COMPARISON OF GROUPS I AND II IN TERMS OF
DIFFERENCES BETWEEN VALUES
FOR LAST AND FIRST EXPRESSED RATE CHOICE

Difference of Last Expressed Rate Choice Minus the First Expressed Choice	Group	Group
	I	II
Positive	27%	31%
Zero	28%	22%
Negative	45%	47%

Honor company ratings failed to reflect sufficient valuability to be useful for the purposes of this study.

Other Evidences

As a support to the school, night school instruction is available. For some students it is mandatory; i.e., those that failed a preliminary mathematics test. Others may attend if they need additional instruction. One session is usually held for all four current classes, both EOG and MOG. Attendance ranges from about 5 to 15 per company. However, some students are required to attend so this was not used as a symptom of attitude.

The correlations among the behavioral evidences of attitude are shown in Table 16.

TABLE 16

CORRELATIONS (PHI COEFFICIENTS) AMONG FOUR
BEHAVIORAL EVIDENCES OF TRAINEE ATTITUDE (N=298)

	1	2	3	4
1. Last Expressed Rate Choice	—	.59	.07	-.28
2. Last Minus First Expressed Rate Choice	.59	—	.03	.24
3. Attitude Questions (4 items)	.07	.03	—	.06
4. Walking Chits	-.28	.24	.06	—

This table offers evidence that there could be a "productive morale" factor in such behavioral data. Coupled with the dif-

ferences shown for Groups I and II, it points to the need to develop such kinds of evidence symptomatic of trainee attitudes relative to policies and practices within their training programs.

VII

CORRELATION AMONG THE TEST SCORES

Table 17 shows the intercorrelations among the factors of the Chicago Primary Mental Abilities Test and the several Marine Corps tests. In this table the correlation coefficients for Group I are shown below the diagonal and for Group II above the diagonal.

The correlations among the Primary Mental Ability factors are lower, as could be expected, than among the Marine Corps measures. The Primary Mental Abilities factors were constructed so as to have maximum independence (or lowest possible intercorrelations among them), while the GCT was constructed with little regard to its factorial organization.

The relation between the coefficients of Groups I and II for the Primary Mental Abilities Test is 0.67 (15 pairs of observations). The relationship between the coefficients between Groups I and II for the Marine Corps tests was higher, a correlation of 0.96. Such differences in consistency of intercorrelation of the two samples could indicate a lower reliability of the Primary Mental Ability Tests than for the sub-tests of the GCT. On the other hand, unless there is some lack of correlation between the two samples, the study cannot reflect differences in the relations of the aptitudes of trainees and the method of training.

Correlations between the Primary Mental Abilities and the Marine Corps tests are of interest. The highest of the correlations is found between factor V: Verbal and the Reading and Vocabulary sub-test of the GCT. There is some similarity, but not identity, between factor N: Number and the GCT Arithmetic Computation, and between factor S: Space and the GCT Pattern Analysis.

TABLE 17

INTERCORRELATIONS AMONG PRIMARY MENTAL
ABILITIES TEST AND MARINE CORPS TESTS

(Note: Group I, below the diagonal, N = 169)
(Note: Group II, above the diagonal, N = 158)

Test	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
PMA: N: Number	(1)	---	201	245	212	273	286	184	337	321	079	231	-074
PMA: V: Verbal	(2)	355	---	271	459	310	258	695	302	425	270	510	237
PMA: S: Space	(3)	205	294	---	176	420	195	221	340	422	485	518	345
PMA: W: Word Fluency	(4)	252	565	322	---	289	402	243	160	222	066	223	-050
PMA: R: Reasoning	(5)	363	513	503	394	---	229	310	384	530	381	517	272
PMA: M: Rote Memory	(6)	192	142	170	090	164	---	206	256	263	110	251	-021
GCT: Read. & Vocab.	(7)	342	663	370	486	492	135	---	393	480	272	612	263
GCT: Arith. Comput.	(8)	552	477	396	331	485	200	462	---	603	407	689	344
GCT: Arith. Reason.	(9)	414	511	325	273	468	190	546	711	---	498	742	367
GCT: Pattern Anal.	(10)	094	094	571	271	471	239	363	458	449	---	761	490
GCT: Total	(11)	401	611	536	426	572	262	684	797	808	745	---	516
Mech. Comprehension	(12)	066	369	430	105	309	127	250	369	403	567	543	---

*Decimal points omitted

VIII

RESULTS

Correlations of Tests Scores and Examination Grades

The coefficients of correlation between the aptitude test scores and the Phase III course examination grades are shown in Table 18 on the following page. The table shows the correlations for both Groups I and II as pairs of coefficients. The correlation for Group I: Rote Method is the upper member of the pair, and just below it is the correlation for the Group II: Key Method.

There is considerable similarity of the correlations of the aptitude test scores and the examination scores for the two methods of instruction. The aptitude test scores as a whole correlated higher with the examination grades for Group I than for Group II. The arithmetic mean of the correlations of the Chicago Primary Mental Abilities Tests with the Phase III examination scores is $\bar{r} = .377$ for Group I, and $\bar{r} = .247$ for Group II. The correlation of the Marine Corps Aptitude Test scores with the Phase III course marks shows a similar difference between Groups I and II ($\bar{r} = .528$ and $\bar{r} = .419$ respectively.) These differences are also reflected by the correlations of the aptitude tests with the centroid factor F_1 . F_1 is the variance common to all of the Phase III course examinations. For each of the tests, the Group I correlation is higher than that for Group II.

In comparing the correlation coefficients for the two methods of instruction, it may be noted that for the Primary Mental Abilities Tests the 5% level confidence requires that the two correlation coefficients differ* by at least .120, and for the 1% level, .160. For the GCT sub-tests, GCT Total and the Navy Mechanical Comprehension to be at the 5% level of confidence, the difference must be at least .133, and for the 1% level, the difference must be at least .175. In Table 18, pairs of correlation coefficients which differ at the 5% level are marked with one star and those significant at the 1% level are marked with a double star.

*The standard error of the difference for comparing the two correlations was computed, using the standard error of r from a universe having a mean of $r=0$, and observing the correlation between the r 's of the two groups to be .69. The correlation between the r 's of the two groups for the Marine Corps Aptitude Tests was .62.

TABLE 18

CORRELATIONS OF APTITUDE TEST SCORES
WITH MARKS IN COURSE EXAMINATIONS(Notes: Group I, N = 169)
(Notes: Group II, N = 158)

Test or Factor	Math I	Math II	Algebra I	Algebra II	Elect I	Elect II	Final Av.	PLQ	Postor Leading P.
MA: N: Number	901 204*	411 224**	300 603**	311 143**	177 109	182 003**	322 305	134 -035**	390 198
MA: V: Verbal	565 417*	567 387**	458 405**	584 365**	518 257**	385 306	614 403**	384 198**	692 498
MA: S: Space	360 315	419 286*	392 356	366 289	312 172*	253 315	444 376	198 057	440 387
MA: W: Word Fluency	390 160**	407 259**	419 161**	437 177**	302 164*	185 119	409 184**	206 -037*	447 229
MA: R: Reasoning	387 405**	614 394**	551 307**	530 382**	386 228*	333 290	655 387**	150 101**	626 446
MA: M: Memory	219 294	257 264	178 139	173 157	184 142	186 070	260 232	155 100	251 235
UAT: Read. & Vocab.	506 322**	463 302**	601 383**	564 353**	376 315**	407 300	563 355**	229 091*	633 144
UAT: Arith. Comput.	651 530	610 518	511 465	564 408*	445 318	395 295	599 527	190 149	665 569
UAT: Arith. Reason.	622 522	600 379	556 434	538 469	519 394	478 423	632 614	198 195	706 637
UAT: Pattern Anal.	509 356*	528 451	450 479	466 383	470 325*	408 371	570 484	161 046	591 535
UAT: Total	710 495**	699 592	652 612	657 538	626 453**	533 445	747 631	238 145	811 699
Mech. Compreh.	404 341	430 374	458 443	435 431	467 230**	462 408	572 454	368 328	564 500

Decimals omitted.

* Difference of the two correlations significant at the 5% level.

** Difference of the two correlations significant at the 1% level.

One might next look at the differences in correlations between the aptitude tests and Phase III examinations for the two methods of presentation from the point of view of the examinations.

Electricity I shows the greatest differences between its correlations in Groups I and II with aptitude test scores. The tests showing significant differences are the V: Verbal, R: Reasoning, the GCT Reading and Vocabulary, the GCT Total Score, and the Navy Mechanical Comprehension Test. With the exception of the last named test, these are tests which are central to the usual concepts of intelligence. The N: Number factor, Arithmetic Reasoning, and Arithmetic Computation each show quite similar correlations for Groups I and II.

Mechanics I and II show similar results. The greatest discriminations for both of these evidences of learning are shown in the correlation with the tests V: Verbal, W: Word Fluency, R: Reasoning, and with the GCT Reading and Reasoning.

The Mathematics I and II examinations show much the same pattern of differences in correlation with the aptitude tests as do the Mechanics examinations.

Partial Regression of Aptitude Tests on Examination Grades

Another view of the results presented in Table 18 is shown in Tables 19 and 20.

Table 19 shows the partial regression coefficients for each of the 12 aptitude tests on each of the six Phase III examinations, the Final Average for the course, the Performance Achievement Quotient (PAQ) and Factor F, successively as dependent or criterion variables. The partial regression coefficient is an indicator of the unique contribution or net relationship of the test score to the dependent variable. The test showing the greatest discrimination between Group I and Group II for Phase III examinations is factor R: Reasoning. Three other factors of the Primary Mental Abilities Test, Number, Verbal and Word Fluency, along with the Reading and Vocabulary sub-test of the GCT are next in order in their discrimination. In each case, the regression coefficient between the aptitude test score and examination mark is higher for Group I: Note than for Group II: Why. It is interesting to note that only the Arithmetic Reasoning of the GCT showed no discrimination at the levels marked. Of the Primary Mental Abilities factors, neither the Space factor nor the Memory factor showed much difference in their correlations with the Phase III examinations for the two methods of presentation. Correspondingly, the GCT Pattern Analysis also showed little difference.

The factor scores of the Primary Mental Abilities seem to be more useful for the purpose of this study than do the Marine Corps Aptitude Tests. On the other hand, the Marine Corps Aptitude Tests are

TABLE 19

PARTIAL REGRESSION COEFFICIENTS OF THE APTITUDE TESTS ON EACH OF THE PHASE III EXAMINATIONS,
THE FINAL COURSE AVERAGE, AND THE PERFORMANCE ACHIEVEMENT QUOTIENT (PAQ)

Test	Group	Math I	Math II	Mech I	Mech II	Elect I	Elect II	Final Av.	PAQ	P ₁
PMI: N: Number	Rate Way	196 107	092 089	-028 -145	-029 018	-136 -006	-028 -137	-023 003	-121 -072	013 -004
PMI: V: Verbal	Rate Way	144 349*	126 211	256 378	118 130	103 -033	029 095	119 296	189 312	193 244
PMI: S: Space	Rate Way	-046 018	-030 -130	-001 069	-078 -036	-071 -078	-090 113*	-056 009	-133 -053	-077 -032
PMI: M: Word Fluency	Rate Way	076 -148*	074 030	037 -048	123 034	023 079	-044 032	073 -039	057 -171	052 008
PMI: R: Reasoning	Rate Way	174 146	156 150*	162 -017	101 079	-010 -028	034 031	228 020	245 021	166 052
PMI: K: Memory	Rate Way	021 122	086 071	026 -005	021 -021	060 007	068 -031	083 060	105 137	055 038
GCT: Read. & Vocab.	Rate Way	000 -128	-147 -231	168 -164*	183 -049*	362 103*	157 066	066 -296	-013 -216	112 -091
GCT: Arith. Comput.	Rate Way	129 334*	062 138	052 172	211 047	063 029	-005 042	067 143	301 043	112 162
GCT: Arith. Reason.	Rate Way	136 224*	105 247	110 -004	033 105	318 174	155 246	114 289	-015 165	176 209
GCT: Pattern Anal.	Rate Way	285 -018*	090 081	038 026	059 -047	162 093	030 141	031 057	-097 -217	121 082
GCT: Total	Rate Way	-246 -154	229 240	011 346*	013 296*	-168 200*	154 -108*	155 185	-218 -007	065 177
Mech. Comprehension	Rate Way	055 060	065 070	166 107	206 197	245 007*	277 291	295 125	379 339	210 199
R	Rate Way	79 67	78 68	76 69	74 59	72 46	60 54	83 70	52 45	88 77

Decimals omitted.

* Difference at least .200.

more effective than the Primary Mental Abilities tests for the purposes for which these tests were designed. Such evidence opens the door to these conclusions:

1. Most of the aptitude tests used in this experiment consistently correlate more highly with marks on the course examinations of those groups taught by the rote method than for those taught by the why method.
2. Learning by the rote method is more closely related to such aptitudes as Number ability (N), Verbal ability (V), Reasoning (R), and Word fluency (W) than is learning by a why or explanation method.
3. The performance on the examinations may be more consistent and uniform when course materials have been presented to be learned by rote methods, thus permitting higher correlations. The course examinations may be a more adequate measure of material learned by rote methods rather than that approached by a more questioning attitude, since this is apparently the method usually followed and the examination questions have been validated within essentially that method.

The Primary Mental Abilities Test, chosen because of its presumed sensitivity to the situation, showed 21 of its 36 pairs of correlations to differ at the 1% or 5% level of significance. Sixteen of the 21 were at the 1% level. The Marine Corps Aptitude Tests showed 11 of its 36 pairs of correlations with Phase III examinations to differ in a statistically significant amount (7 at the 1% level and 4 at the 5% level).

Table 19 also shows the multiple correlation coefficient of the twelve tests for each dependent variable for Group I and for Group II. These coefficients (R) shown at the bottom of the table are the correlations of the examination with the optimally weighted combination of all twelve aptitude test scores. For example, the twelve tests combined at the best linear weights correlate $R = .79$ with the Math I examination for Group I and $R = .67$ for Group II.

For all but one of the Phase III examinations the value of R is greater for Group I, taught by the rote method. On the average, this difference in the multiple correlations is .15. Nineteen of the 72 regression coefficients of the Marine Corps Aptitude Tests in the 6 Phase III examination grades are .200 or greater, while only 5 of the 72 Primary Mental Abilities factors have regression coefficients as great. The Marine Corps Aptitude Tests carry the greater part of the greater portion of the predictive load.

Only three of the aptitude tests show consistent results for the majority of the examinations. Factor V: Verbal and the GCT total score show slightly higher regression coefficients for Group II, while the Navy Mechanical Comprehension test shows the reverse.

Each pair of such coefficients of Group I and Group II showing a difference of at least .200 has been starred in Table 19. This difference was chosen arbitrarily, since the computation of a critical ratio for partial regression coefficients is a somewhat lengthy process. For the purposes of this study, the selection of some such arbitrary value will serve. Four such differences are shown for the Primary Mental Abilities factors:

- V: Verbal (Math I)
- S: Space (Electricity II)
- W: Word Fluency (Math I)
- R: Reasoning (Math II)

In the first two factors, the coefficients for Group II: Why was greater than for Group I: Rote. For the last two factors the reverse was true.

Factor V shows consistent direction of difference favoring the why method in all Phase III examinations except Electricity I, and Factor R: Reasoning shows a larger coefficient for Group I than for Group II.

This points to the conclusion that while Factor V: Verbal is central to success in Phase III, whether taught by one or the other method, trainees with higher V scores tend to get better marks when taught by the why method, and those with higher R scores tend to get better marks when taught by the rote method. The principles indicated above are supported by the partial regression coefficients for the Marine Corps Test shown in the same table. The GCT total score shows larger coefficients for Group II: Why, than for Group I: Rote, paralleling the evidence shown by factor V: Verbal. The GCT Reading and Vocabulary test shows more nearly the same results as factor R: Reasoning. The zero order intercorrelations of the tests indicate, however, that this test is a little more similar to the V: Verbal factor than to the R: Reasoning factor. Much of the possible net regression of the Primary Mental Abilities factors on the examination marks may be overshadowed when using the Marine Corps Aptitude tests in the same regression equation.

Table 20 shows the partial regression coefficients of the Primary Mental Abilities factors on each of the Phase III examinations, on the final course average, the Performance Achievement Quotient, and on the F₁ centroid factor. In general, the regression coefficients for the rote group and for the why group are quite similar. In Table 20 the regression coefficients for N: Number are greater for

TABLE 20

PARTIAL REGRESSION COEFFICIENTS FOR THE CHICAGO PRIMARY MENTAL ABILITIES
TESTS ON EACH OF THE PHASE III EXAMINATIONS, THE FINAL COURSE
AVERAGE AND THE PERFORMANCE ACHIEVEMENT QUOTIENT (PAQ)

PMA Factor	Group	Math I	Math II	Mech I	Mech II	Elect. I	Elect. II	Final Av.	PAQ	FI
N-Number	Rote	263	138	007	045	-081	003	018	-060	077
	Why	182	063	-154	-011	000	-161	009	-101	-018
V-Verbal	Rote	278	277	493	358	453	335	387	303	470
	Why	327	260	449	282	182	243	297	256	391
S-Space	Rote	053	110	122	093	132	110	150	-042	130
	Why	060	064	246	123	055	243	200	007	177
W-Word Fluency	Rote	025	039	006	095	-029	-105	001	-036	003
	Why	-174	-025	-105	-043	012	-045	-091	-207	-085
R-Reasoning	Rote	301	332	226	238	112	118	307	236	274
	Why	246	249	130	236	134	173	212	080	262
M-Memory	Rote	068	114	048	050	097	109	126	095	102
	Why	160	119	032	006	048	-016	100	128	079
R	Rote	71	72	71	66	56	44	73	44	72
	Why	58	51	57	46	31	44	53	29	61

Group I: Rote than for Group II: Why. Factor V: Verbal is the most important aptitude shown in the table. Its regression coefficients are the largest of any of the aptitude factors. Its differences for Groups I and II are, in general, in favor of Group I.

If one considers the F_1 factor as the best overall evidence for learning in Phase III, then, regardless of method of instruction, those having higher scores in V: Verbal and W: Word Fluency do best on the examinations. The evidence from Table 20 leads to one of two conclusions:

1. The two methods of teaching do not draw upon markedly different patterns of aptitudes.
2. The two methods of instructional performance were not significantly different.

Phase I, Unit 4a: Aircraft Familiarization

Further evidence regarding the hypothesis of the study is offered in the results from the teaching of Phase I, Unit 4a: Aircraft Familiarization by the two methods. Only classes 28 (rote) and 29 (why) were involved in this part of the experiment.

The data are shown in Table 21. The correlation coefficients and regression coefficients for Group I: Rote are shown below and to the left of the diagonal and the corresponding data for Group II: Why are above and to the right of the diagonal. Intercorrelations among the aptitude tests were used in computing the partial regression coefficients, but need not be examined here for other purposes.

One might first compare the correlations of each aptitude test with the examination in Aircraft Familiarization for the two methods of instruction. The largest differences which show a larger correlation coefficient for Group I: Rote than for Group II: Why are the following aptitude tests: S: Space, W: Word Fluency, M: Memory, Arithmetic Comprehension, Arithmetic Reasoning, and GCT Total Score. None of the aptitude tests showed correlations with the examination noticeably greater for Group II: Why than for Group I: Rote.

The comparison of the partial regression coefficients for Group I and Group II, shows that W: Word Fluency is much more important for the rote method than for the why method and that while M: Memory carries practically no importance for rote learning, it has a negative weight for the why method.

The V: Verbal factor is important to both methods of learning, but those with a high score on W: Word Fluency learn better by the rote method, and those with a low score on M: Memory do better on the why method. When the Marine Corps Aptitude Tests are also taken

TABLE 21

INTERCORRELATIONS AMONG TESTS FOR CLASSES 28 AND 29 AND
CORRELATIONS OF APTITUDE TEST SCORES WITH EXAMINATION
SCORES FOR AIRCRAFT FAMILIARIZATION (I, 4a)

(Class 29 above the diagonal, N = 71)
(Class 28 below the diagonal, N = 86)

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	Exam	Par. Reg.	Par. Reg.
PMA: N: Number	—	079	230	157	235	270	221	337	386	-007	219	-055	-252	-278	-248
PMA: V: Verbal	249	—	260	392	301	207	705	346	438	414	582	357	447	172	491
PMA: S: Space	075	172	—	224	299	134	103	111	300	343	368	343	-012	-130	-102
PMA: W: Word Flec.	092	467	302	—	313	503	240	288	373	169	335	077	065	038	-013
PMA: R: Reasoning	323	407	507	427	—	179	234	274	476	292	431	221	187	116	170
PMA: M: Memory	167	226	076	102	143	—	172	300	366	115	215	-129	-166	-238	-211
GCT: Read. & Vocab.	240	532	215	357	342	067	—	337	406	257	507	181	422	356	—
GCT: Arith. Comp.	441	460	320	216	460	195	354	—	562	276	597	178	096	042	—
GCT: Arith. Reason.	335	483	168	082	353	155	469	638	—	403	688	306	186	151	—
GCT: Pattern Anal.	-055	307	520	406	490	140	237	349	353	—	818	523	352	357	—
GCT: Total	237	567	457	389	523	212	535	760	764	727	—	497	336	-318	—
Mech. Comprehension	-039	363	405	159	196	111	077	355	340	547	504	—	335	099	—
Exam	-051	536	269	522	360	149	354	335	348	463	504	456	—	—	—
Partial Regression	-246	247	-009	339	037	044	062	151	222	189	-244	108	—	—	—
Partial Regression	-223	390	074	287	105	054	—	—	—	—	—	—	—	—	—

Decimals omitted.

into account, it may be concluded that, other things being equal, those trainees with higher scores on the Reading and Vocabulary test and on the Pattern Analysis test learn better when taught by the why method.

The patterns shown by the above are illustrated in Table 22.

TABLE 22

AVERAGE EXAMINATION MARKS
FOR AIRCRAFT FAMILIARIZATION

Test Score Pattern	Class 28	Class 29
	Rote	Why
Had above average score on M: Memory	80	78
Had below average score on M: Memory	78	83
Had above average score on W: Word Fluency	86	80
Had below average score on W: Word Fluency	70	80

The data in this table shows that those trainees with below average scores on M: Memory receive higher average examination grades when taught by the why method rather than the rote method, and how those with above average scores on W: Word Fluency receive better examination marks when taught by the rote method rather than by the why method.

The Navy Interest Test

Correlations of the Navy Interest Test scores with each of the course examinations for Group I: Rote and for Group II: Why are shown in Table 23. In each pair of correlations, the upper number is for the Group I: Rote sample and the lower one for the Group II: Why sample.

A correlation coefficient for Group I: Rote must be at least .151 to be considered more than just a random variation from a true value of zero (5% level of confidence). The corresponding value for Group II: Why is .155. The coefficients that are significant at the 5% level are followed by one dot and those large enough to be significant at the 1% level are marked with two dots. Three Interest scores show fairly consistent and significant positive correlations with the examination marks in the rote sample:

TABLE 23

CORRELATIONS OF NAVY INTEREST TEST SCORES
WITH SCORES IN COMBAT REGISTRATION(Data: Group I, N = 149)
(Data: Group II, N = 156)

Functional Interest Key	Math I	Math II	Math I	Math II	Math I	Math II	Fire-Fight.	Alt. Fuel.	Tech. Pub.	Final A/c.	PG
(AP) Aviation Technician (Electronics)	063	050	197	190	110	205	104	063	006	082	161
(AP) Aviation Mate (Mechanical)	046	069	006	019	199	001	-006	075	042	051	099
(CP) Construction Apprentices	192	277	197	286	203	110	199	130	139	253	095
(AM) Aviation Structural Mechanist	-015	056	037	109	117	116	197	113	215	129	144
(AO) Aviation Ordnanceman	-046	-114	-048	-103	004	019	-047	-107	-044	-081	080
(BM) Boatman's Mate	046	006	092	-007	-003	-063	001	056	056	-075	-015
(CB) Communications	147	146	117	107	177	076	165	197	146	216	075
(CT) Communications Technician	064	046	133	006	-080	141	213	042	114	111	046
(DC) Damage Controlman	151	191	176	170	204	103	106	208	117	223	182
(EM) Electrician's Mate	000	067	157	061	-079	063	140	101	001	074	-040
(ET) Electronics Technician	032	-063	-146	-177	-117	-148	-175	-113	-079	-134	-223
(FC) Fire Controlman	-116	-065	-097	-190	-237	-138	011	-046	-139	-180	-135
(GM) Gunner's Mate	-199	-201	-172	-235	-176	-143	-186	-163	-118	-214	-089
	-171	-203	-165	-253	-103	-235	-183	-183	-165	-207	-205
	-054	-109	089	-043	-066	-028	001	-061	-046	-028	-022
	027	-049	-080	-108	024	-081	-153	-010	-026	-092	-033
	079	085	037	054	031	-040	-006	056	033	019	-009
	-050	-019	-080	-044	-110	-058	047	-099	-102	-067	-126
	018	-060	010	007	-020	125	073	-023	-063	003	069
	-009	-037	-097	-006	059	-006	-049	-015	-091	011	048
	021	-015	078	099	-016	134	023	002	-006	013	088
	139	139	061	110	126	113	-050	094	147	155	183
	-019	-012	056	001	-021	139	022	-028	-059	-013	092
	012	-010	-078	-010	130	018	-126	-037	-033	014	098
	137	046	126	076	061	089	141	126	064	130	-024
	-084	-051	-081	-043	-039	-030	046	-078	-004	008	-027

TABLE 23

CORRELATIONS OF NAVY INTEREST TEST SCORES
WITH SCORES IN COURSE EXAMINATIONS

(Cont'd.)

Verbal Interest Test	Math I	Math II	Math I	Math II	Elect. I	Elect. II	Phys. I	Phys. II	Final Avg.	Final Avg.
(18) Hospital Coursework	-.222**	-.172**	-.047	-.108	.022	-.017	-.041	-.008	-.099	-.049
	-.044	.077	-.011	-.018	.073	-.108	-.153	-.114	-.110	-.034
(19) Mathematics	-.167**	-.167**	.086	-.112	-.078	-.010	-.077	-.073	-.141	-.023
	.092	.077	.023	.000	.048	.075	-.008	.124	.077	.136
(20) Stenokeyper	-.141**	-.154*	-.267**	-.241**	-.298**	-.262**	-.311**	-.233**	-.251**	-.186*
	.154*	.073	-.071	-.076	-.020**	-.136*	-.077	-.139	-.077	-.137
(21) Mechanical	.187**	.177*	.214**	.142	.109	.257**	.187*	.114	.187*	.113
	-.044	-.013	-.087**	.042	.044	.046*	.144	.025	.067	.034
(22) Medical Services	-.114	-.129	.100	.049	.100	.041	.076	.066	-.022	.073
	.009	.017	-.080	.017	.020	-.063	-.100	-.046	-.013	-.008
(23) Clerical	-.217**	-.016**	-.272**	-.243**	-.258**	-.118**	-.247**	-.216**	-.242**	-.162**
	.157*	.133	.083	.077	.128	.073	-.046	.083	.090	.092
(24) Electronics	-.093	-.112	-.089	-.042	-.074	.091	.040	-.068	-.067	.026
	-.046	-.076	-.143	-.067	.013	-.071	-.114	.008	-.022	.044
(25) Food Services	-.040	-.070	-.083	-.101	-.073	-.070	-.049	-.063	-.103	-.115
	-.076	-.047	-.046	-.076	-.074	.073	-.122	-.126	-.103	-.154
(26) Bookkeeping	.102	.157*	-.041	.076	.015	-.017	-.046	.064	.102	.074
	.003	.077	.046	-.009	-.128	.081	.130	-.044	.006	-.117
(27) Physical Activities	-.174**	-.120**	-.102	-.086	-.077	-.146	-.080	-.130*	-.144	-.067
	.113*	.126*	.116*	.107*	.102	.000	-.013	.051	.026	.014
(28) Class Room Activities	-.159**	-.271**	-.074**	-.236**	-.279**	-.118**	-.236**	-.274**	-.332**	-.188
	.074*	.011	-.016**	.049**	.004	.002	-.084	-.048*	-.007**	.050*
(29) Physical Effort	.072	.024	.042	-.022	.042	.013	.023	.032	.015	-.038
	-.102*	-.127	-.113	-.122	-.176*	-.207**	.021	-.115	-.067	-.072

* A difference between the degree of correlation shown and zero correlation is significant at the 5% level of confidence.
 ** A difference between the degree of correlation shown and zero correlation is significant at the 1% level of confidence.
 *** A difference between the degree of correlation shown and zero correlation is significant at the 0.1% level of confidence.
 **** A difference between the degree of correlation shown and zero correlation is significant at the 0.01% level of confidence.
 ***** A difference between the degree of correlation shown and zero correlation is significant at the 0.001% level of confidence.
 Decimals omitted.

AD - Aviation Mate (Mechanical)
AO - Aviation Ordnanceman
H1 - Mechanical Group

Three Interest scores show fairly consistent and significant negative correlations with the examination marks in the rote sample:

SK - Storekeeper
H3 - Clerical Group
H8 - Clean Hands Activities

Only one key shows any consistent and significant correlation with examination marks for the why group. CS - Commissaryman shows negative correlations for both the why and the rote groups.

Such correlations do generally demonstrate validity for the Navy Interest Test. However, the lack of correlation with examinations for the why group is not explained. For the purposes of this study, the differences in value within each pair of coefficients are the pertinent data.

Each pair of Group I and Group II correlation coefficients which differ at the 5% level of significance (.177) is marked with a star and if significant at the 1% level (.231), two stars. Of the 150 pairs of correlations with Phase III examinations, 38 are so marked. Two of the Interest Scores, H3 - Office and Clerical and H7 - Verbal Activities show significant and consistent differences for each of the six comparisons. SK - Storekeeper shows significant differences for five of the six comparisons. H8 - Clean Hands Activities shows a similar but not as significant relationship. These four Interest Scores correlate negatively with the examination scores for the Group I: Rote sample and about zero in the Group II: Why sample. Such data could reflect a dislike for the MOG rote activities, but did not do so in Group II. It is possible that these do reflect attitudes negatively related to one method of instruction and unrelated to another.

Four other Interest Scores show consistent differences in their correlations with the Phase III examination scores.

AD - Aviation Mate (Mechanical)
AO - Aviation Ordnanceman
GM - Gunner's Mate
H1 - Mechanical

While few of the differences are large enough to be individually significant, the trend suggested may have some real importance; that the rote method of instruction is better responded to by those whose interests are typical of rates open to MOG trainees, but the why method of instruction is responded to positively or negatively by trainees whose interests are not those of MOG rates.

IX

SUMMARY

The data just presented lead to the following conclusions:

1. When the methods of instruction differ consistently, the patterns of aptitudes required for success do show reasonable and acceptable differences.
2. When the methods of instruction do not differ clearly and consistently, the patterns of aptitudes do not and cannot reflect sensible differences.
3. Within the Mechanical Occupations Group, those with the interests characteristic of mechanical rates, learn better by the rote method. This could be a reflection of the fact that the trainees included in the study were assigned to be trained for mechanical occupations. Those whose interests coincided did better when trained in a manner typical for such kinds of occupations.

More specifically, the data presented showed the following:

1. The correlations of aptitude test scores with examination marks were higher for the rote presentation than for the why presentation. This is probably due to the examinations used, rather than to differences in consistency of performance.
2. The average examination marks were higher for the why presentation than for the rote presentation. It is possible that:
 - a. Trainees learn more when the why method is used.
 - b. Trainees score higher on examinations when the same examinations were used by immediately preceding classes. This factor could be operative also in reducing or alternating the correlations for the why presentation relative to the rote presentation.
3. Over all, the Marine Corps Aptitude Tests correlated higher with examination marks than did the Primary Mental Ability Tests.

4. In general, the Primary Mental Ability Tests were more sensitive to differences in the methods of instruction than were the Marine Corps Aptitude Tests.
5. The partial regression coefficients of the aptitude tests on course examinations in Phase III show that:
 - a. V: Verbal factor and GCT total score contribute to the examination scores for both methods of instruction, but more to the why method.
 - b. R: Reasoning and GCT Reading and Vocabulary Test contribute more to the examination scores for the rote presentation.
6. Considering the partial regression coefficients for the Primary Mental Abilities Tests without the Marine Corps Aptitude Tests:
 - a. N: Number contributes more to the examination scores for the rote presentation.
 - b. V: Verbal and R: Reasoning are important to both modes of presentation.
7. In so far as there is not a clear-cut difference in aptitude patterns for learning by the rote presentation and by the why presentation -- Phase III -- is likely a reflection of a lack of a clear-cut difference in the two modes of presentation as actually practiced. The classroom observer on several occasions reported that he could not see sharp differences in the teaching when following the rote method and when following the why method in Phase III, although his observations did show differences in the "kys per hour" shown in Table 8. He also reported that differences in the instruction by the two methods in Aircraft Familiarization seemed more distinct and clear.
8. The experimental data in Phase I, Unit 4a: Aircraft Familiarization show some clear differences in aptitude patterns required for success in the two instructional methods:
 - a. The correlations of the aptitude test scores with the examination marks in Aircraft Familiarization are higher for the rote presentation than for the why presentation.
 - b. The partial regression coefficients show the following pattern: